

## Human genetics in the Crescent City

Genome-wide association studies are now being carried out for a variety of human traits. Announcements at the annual meeting of the American Society of Human Genetics, held recently in New Orleans, offer a glimpse into the changing landscape surrounding issues such as study design, reporting, analysis and data accessibility.

The American Society of Human Genetics (ASHG) should be congratulated for having just held its annual meeting in New Orleans, honorably fulfilling a commitment to a city still struggling to recover from Hurricane Katrina. A glance at the abstracts confirmed that genome-wide association (GWA) studies are now being carried out for a variety of human traits. Improved genotyping technologies, the HapMap and large consortia of investigators have enabled such studies, which are the obvious next step for human geneticists. Announcements both before and at the meeting provided an opportunity to survey the changing landscape.

In addition to the GWA studies currently underway, the NIH earlier this year announced the Genetic Association Information Network (GAIN). GAIN involves Pfizer, Perlegen and Affymetrix in a collaborative effort with the NIH to carry out GWA studies for several common diseases. In New Orleans, Francis Collins of the National Human Genome Research Institute (NHGRI) and Patrice Milos of Pfizer announced the selection of six peer-reviewed, investigator-initiated proposals for funding. They include GWA studies of psoriasis, attention-deficit hyperactivity disorder, schizophrenia, bipolar disorder, depression and nephropathy in type 1 diabetes. It seems noteworthy that three of the six involve common and devastating psychiatric disorders, where genetic insights that might shed light on the underlying pathophysiology are urgently needed. Genotyping will begin by the end of this year, and data will be made available to other researchers during the first quarter of 2007.

Several aspects of GAIN should make it a standard-setting effort. First, a statistical analysis workshop open to the entire community, to be held at the end of November, will aim to reach a consensus on the most powerful and appropriate methods of data analysis. Having witnessed a lack of consensus in the formal review of genetic association studies submitted to this journal over the last 2–3 years, we urge the widest possible participation, as well as a wide dissemination of guidelines for study design, data reporting and analysis.

These are not the only thorny issues to be resolved. In response to a question about the delay in announcing the recipients of grants under the medical sequencing program of the NHGRI, Collins noted that the proposals had been reviewed but that concerns about data accessibility and privacy issues were still being aired. GAIN clearly faces the same issues, which is why the NIH issued a request for information in August asking for advice and laying out proposals for data submission and access, privacy protection, publication and intellectual property.

In our view, the broad principles laid out in the proposal are reasonable and deserve support in that they promote accessibility of the data to a wide range of researchers while taking sensible precautions to guard against the small chance that such openness could be abused. In regard to

data submission, investigators receiving NIH funds will be asked to submit the protocol, questionnaires, variables measured and any other relevant documentation. These data will be available to everyone. Submission of more detailed coded (de-identified) data covering phenotype, exposure, genotype and pedigree data will be submitted under stricter control. Local institutional review boards would be charged with certifying that the submission to the central repository takes account of informed consent and contains adequate safeguards to protect the identities of the participants. Third-party researchers wanting to use this 'second tier' of data would be granted access by an NIH Data Access Committee, which will ensure that the data are used only for approved uses. If this structure is adopted, we would encourage authors at the time of submission to document the source of the data and to state that they are in full compliance with the provisions of the Data Access Committee. The NIH also proposes a 9-month period following data submission during which the generators of the data would have exclusive rights to publish an analysis. This policy has the virtue of specifying an unambiguous period of time for exclusive publication, after which ethical publication would be available to any researcher. The issue of intellectual property may actually be the least contentious. At a workshop earlier this year to discuss similar policies for The Cancer Genome Atlas, industry representatives were clear that genotype data should be considered pre-competitive, a position echoed by Pfizer's Milos at the ASHG meeting.

What can we expect when the results start pouring in? Notably absent from the above programs is any mention of copy number or other structural variation. Participants at the ASHG meeting noted the importance of simultaneous assessments of SNPs and copy number variants, in aid of the fullest possible picture of common genomic variation. Epigenetic variation—a dark horse in the search for risk factors—will also have to be accounted for, although the integration of such data with genetic variation may be a difficult task.

Preliminary results reported at the meeting suggest that some dramatic recent findings—*CFH* in age-related macular degeneration, and *TCF7L2* in type 2 diabetes—may not be representative of the majority of common risk variants. In a comment on the recent paper by Jonathan Flint and colleagues describing characteristics of mouse quantitative trait loci (*Nat. Genet.* **38**, 879–887; 2006), Thomas Mitchell-Olds noted the "fundamental tragedy of biomedical genetics: that quantitative genetics is true"—meaning that the vast majority of risk variants will have a very small influence on disease susceptibility. Although 'tragedy' is used here with tongue in cheek, it is clear that the separation of true variants of modest effect from false ones remains a substantial, if not insurmountable, challenge. ■